Ann. For. Sci. 61 (2004) 109–115
© INRA, EDP Sciences, 2004
DOI: 10.1051/forest:2004001

Original article

Stem profile equations for young trembling aspen in northern Ontario

Michael T. TER-MIKAELIAN[a]*, Wojciech T. ZAKRZEWSKI[a], G. Blake MACDONALD[a], David H. WEINGARTNER[b]

[a] Ontario Forest Research Institute, 1235 Queen Street East, Sault Ste. Marie, Ontario P6A 2E5 Canada
[b] 118 Country Club Place, Sault Ste. Marie, Ontario P6A 6A3 Canada

(Received 5 February 2002; accepted 18 February 2003)

Abstract – Stem profile equations were developed for young trembling aspen (*Populus tremuloides* Michx.) growing on 6 different sites in northern Ontario. The objective was to develop a model using stem diameter at any height to predict stem diameter at other heights and then compare the model among sites. The 262 sample trees were split into calibration (180) and validation (82) sets. Four stem profile models having from 1 to 3 parameters were tested for parameter significance for each site, and then compared among sites. Data analysis was performed using bootstrap methods. One of the 4 models was not significantly different ($P > 0.278$) among sites. Validation of this model produced mean prediction errors and mean relative prediction errors of less than 0.001 cm and 0.09%, respectively; standard deviations of these errors were 0.15 cm and 4.09%, respectively. The presented method can be used to develop stem profile equations that are applicable to a broad range of growing conditions.

***Populus tremuloides* / stem profile models / stem diameter / bootstrap methods**

Résumé – Équations de profil de tige de jeunes trembles de l'Ontario septentrional. On a mis au point des équations de profil de tige pour des trembles (*Populus tremuloides* Michx.) se développant dans 6 stations différentes du nord de l'Ontario. L'objectif était de créer un modèle utilisant le diamètre pris à n'importe quelle hauteur pour prédire le diamètre à d'autres hauteurs et ensuite pour comparer les modèles selon les sites. Les 262 arbres témoins ont été affectés soit au calibrage (180) soit à la validation (82). Sur quatre modèles comportant de 1 à 3 paramètres, on a d'abord testé le caractère significatif des paramètres dans chaque site, puis procédé à des comparaisons entre sites. Pour l'analyse des données on a utilisé les méthodes « bootstrap ». L'un des 4 modèles ne présentait pas de différence significative ($P > 0,278$) entre sites. La validation de ce modèle inique une prédiction d'erreur moyenne et d'erreur relative moyenne respectivement inférieurs à 0,001 cm et 0,09 %. Les écarts types de ces erreurs étaient de 0,15 cm et 4,09 %. La méthode présentée peut être utilisée pour développer des équations de profil de tige applicables à une vaste gamme de conditions de croissance.

***Populus tremuloides* / diamètre de tige / méthodes « bootstrap »**

1. INTRODUCTION

Estimating individual tree volume with better accuracy has always been an essential component of forest inventories. The most common approach of obtaining these estimates is to use stem profile equations that predict diameter at any given point on the stem from a few easily measured variables [14]. Desirable features of the stem profile models are that they should provide an opportunity to directly estimate height for any stem diameter, and should be capable of being integrated to give a compatible volume function [10, 14]. In the past few decades, forest mensurationists have developed numerous stem profile equations [1, 4–6, 8, 14, 15, 21, 22]. An extensive review of stem profile equations was published in [10].

Tree age is known to have a profound effect on stem profile [9, 10]. The stem of a mature tree can be divided into three sections based on form, namely: the crown, the clear bole, and the butt swell [9]. The profiles of these sections differ, and consequently, are best described by different analytical models. Research efforts to develop a stem profile model that accounts for these differences in section profiles have produced such models as the variable-exponent [8], the variable-form [5, 14], the segmented variable-exponent [4], and the switching model [21].

However, stem profile segmentation is much less pronounced in young trees. These normally have long crowns and strong tapering stems typical of a crown segment of a mature tree [9]. For example, Forslund [6] studied the stem profile of different size classes of aspen trees and suggested that trees

* Corresponding author: michael.termikaelian@mnr.gov.on.ca

begin their life as paraboloids, with differences in section profiles increasing with tree size. Therefore, he suggested that the stem profile of a young tree may be adequately described by a simpler model with fewer parameters than those required for a mature tree. Similar results were obtained by Allen [2] who found the stem profile of small (stem volume less than or equal to 0.5 m^3) Caribbean pine (*Pinus caribaea* Morelet var *hondurensis* Barrett & Golfari) trees to be much closer to a paraboloid form than that of medium and large trees. Consequently, his polynomial model used fewer parameters to describe the stem profile of the small trees [2].

Several additional features should be considered when developing a stem profile model for young trees. First, many stem profile equations require diameter at breast height (DBH) as an input variable [8, 10, 14, 19]. Young trees are often too small to be adequately characterized by DBH, requiring stem diameter to be measured at a lower height for model input. Second, two major sources of data for stem profile modelling are permanent sample plot and stem analysis data. Calibration of stem profile equations involves the use of repeated measurements taken from the same trees (repeated over time and/or at various stem heights). These repeated measurements require the use of specialized regression methods such as the generalized least squares approach [10] or bootstrap methods [3, 7, 11, 16]. The advantage of the bootstrap approach is that it is free of additional assumptions about the model, such as the structure of error terms. It also uses calibration data more efficiently by alternating various stem diameter measurements as predicting and predicted values.

Finally, it is desirable that the stem profile equations be applicable to a wide range of site, stand, and tree characteristics; for a discussion of factors affecting tree form and taper see [9, 10]. Attempts to relate stem profile equations to site- and/or tree-specific factors have been inconsistent. For example, Morris and Forslund [12] found that microsite and climatic variables explained 61.9% and 38.7% of variation in stem taper and shape, respectively, in Forslund's model for jack pine (*Pinus banksiana* Lamb.). Muhairwe et al. [13] found that adding age, site, and/or tree crown class variables to Kozak's [8] model only marginally improved model fit for Douglas fir (*Pseudotsuga menziesii* (Mirb.) Franco), trembling aspen (*Populus tremuloides* Michx.), and western red cedar (*Thuja plicata* Donn). However, these attempts have usually been restricted to testing one a priori chosen model, a limited list of site and/or tree variables, and a specific algebraic way of including these variables in the model. The question of whether the studied trees can be described by a single profile equation remains unanswered.

In our study, we developed a model for young trembling aspen trees, using an approach designed to account for desired features of stem profile models discussed above. The approach was applied to young trembling aspen trees growing on a variety of sites. The specific objectives of the study were:

(a) to develop a number of stem profile models of young trembling aspen trees capable of using stem diameter measured at any height as an input variable,

(b) to compare these models among 6 trembling aspen populations from different growing environments, and

(c) to evaluate and compare stem diameter predictions from these models, using a validation dataset.

2. MATERIALS AND METHODS

2.1. Study sites and data collection

Data for this study were collected during the establishment phase of a research trial designed to examine the effects of spacing on the growth of young trembling aspen. Six study sites were located in pure aspen (over 80% of the basal area) stands in northern Ontario: Timmins (48° 18' N, 81° 18' W), Dryden (49° 53' N, 92° 53' W), Mann Lake (49° 30' N, 85° 32' W), East Hillsport Lake (49° 29' N, 85° 29' W), White Otter Lake (49° 30' N, 85° 32' W), and Flanders Road (49° 23' N, 85° 29' W). All the sites had been previously harvested, and no treatments were applied between stand harvest and study establishment on 5 of the 6 sites. The Mann Lake site was site prepared with shark-finned barrels 4 years prior to study establishment and planted the following spring with white spruce (*Picea glauca* (Moench) Voss.) at 2000 stems/ha; when the study was established the site was dominated by aspen, although its density was lower than on the other 5 sites because of the site preparation (see Tab. I). The study was established in 1979-1981 using a randomized complete block design, with 6 thinning treatments per block replicated 4 times on each site. Additional information on the study design and site characteristics is provided in [18].

Two trees from each of the 6 plots were randomly selected and total tree height (H, m), DBH (cm), and stem diameter at 0.3 total tree height (D03, cm), were measured, for a total of 288 trees (2 trees per plot × 6 plots per block × 4 blocks per site × 6 sites). Prior to thinning treatments, the following stand variables were measured for each sample tree in a 2 × 2 m plot centred around the tree: total number of stems taller than or equal to 1 m, total number of stems shorter than 1 m, and total basal area (using a wedge prism).

Excluded from the analysis were 8 trees with missing measurements (e.g., trees shorter than 1.3 m), 9 trees with 2 diameters measured at stem locations within a few centimeters of each other (i.e., total tree height of about 4.35 m), and 9 trees for which the 2 stem diameter measurements indicated an inverted cone shape, reducing total sample size to 262 (Tab. I). For each site, 30 trees (65% to 73% of the sample size) were randomly selected for model fitting, while the rest were used to validate the fitted models.

2.2. Equation development

The objective of the analysis was to develop a model that predicts tree stem diameter D_2 (cm) measured at height h_2 (m) from the stem diameter D_1 (cm) measured at height h_1 (m). The tests included the following models, presented here in a difference form.

(1) Zakrzewski's model [22]

$$D_2 = D_1 \sqrt{\frac{(a-h_1)(z_2^2 + bz_2^3 + cz_2^4)}{(a-h_2)(z_1^2 + bz_1^3 + cz_1^4)}}. \quad (1)$$

Here, D_1, D_2, h_1, h_2 are tree stem diameters and heights as defined above, H is the total tree height (m), z_i (i = 1,2) is an auxiliary variable defined as $z_i = 1 - h_i / H$, and a, b, c are model parameters.

(2) Ormerod's model [15]

$$D_2 = D_1 \left(\frac{z_2}{z_1}\right)^a. \quad (2)$$

(3) Forslund's model [6]

$$D_2 = D_1 \left[\frac{1 - \left(\frac{h_2}{H}\right)^b}{1 - \left(\frac{h_1}{H}\right)^b}\right]^a. \quad (3)$$

Table I. General statistics for variables measured in 6 study sites used in the development of stem profile equations for young trembling aspen.

Site location		Timmins (Site 1)	Dryden (Site 2)	Mann Lake (Site 3)	East Hillsport Lake (Site 4)	White Otter Lake (Site 5)	Flanders Road (Site 6)
Age (year)		5	5	5	10	15	22
Mean stem density of trees taller than 1 m (trees/ha)		52 800	50 700	15 100	27 500	52 000	28 700
Mean stem density of trees shorter than 1 m (trees/ha)		23 300	24 300	20 900	6 000	14 800	45 400
Basal area (m^2/ha)		3.9	7.25	0.57	13.67	14.47	24.82
Number of sample trees		41	44	46	43	43	45
Total height (m)	Min	2.25	2.49	1.36	4.03	4.15	6.41
	Max	5.05	5.05	3.5	6.24	8.69	13.11
	Mean	3.54	3.46	2.22	5.46	5.99	9.91
	St.dev.	0.63	0.57	0.43	0.5	1.03	1.52
DBH (cm)	Min	1.1	1.1	0.1	2	2.2	3.3
	Max	3.3	3.3	2.3	4.7	5.8	11.2
	Mean	2.27	1.98	1.21	3.46	3.93	7.48
	St.dev.	0.48	0.42	0.51	0.59	0.86	1.84
Diameter at 0.3 stem height, D03 (cm)	Min	1.7	1.4	1.1	2.1	2.5	3
	Max	3.4	3.2	2.9	4.3	5.1	10.2
	Mean	2.47	2.12	1.91	3.26	3.63	6.67
	St.dev.	0.37	0.34	0.4	0.5	0.69	1.61

Note that model (3) becomes model (2) when $b = 1$.

(4) Alemdag's model [1]

$$D_2 = D_1\left[\frac{\exp(az_2^b) - 1}{\exp(az_1^b) - 1}\right]^c. \quad (4)$$

Equation (1) is derived from a model for cross-sectional area that can describe stem profiles with up to two inflection points [22]. Equation (2) can describe stem profiles of neiloidal, conical, paraconical, or paraboloidal form depending on the value of parameter a. Equation (3) presents a generalized case of model (2) that can generate stem profiles with one inflection point. Finally, model (4) does not assume stem profiles of any specific geometric form, but is based on the assumption that stems of trees of different size are proportionally similar in all dimensions. For detailed discussion of model properties and figures of stem profiles generated by each model see the original papers [1, 6, 15, 22].

The selected models (1)–(4) were fitted to the data from young aspen trees and compared among sites using the bootstrapping technique described below. To fit a model to a given set of sample data, model parameters were estimated using a nonlinear regression. The loss function was defined as the sum of squared prediction errors (residuals), and its global minimum was sought using a conjugate gradient method [17]. The method assumed symmetric residuals with the mean equal to zero; residuals were tested for homogeneity for a subset of data samples.

The models (1)–(4) were first tested for parameter significance. Parameters of model (1) were tested for significant difference from zero, while parameters of models (2)–(4) were tested for significant difference from one. For each model, the tests were performed separately for each site using a randomization technique [3, 16] as follows:

(1) To form a "randomized" sample, for each tree from a given site, one stem diameter (either DBH or D03) was randomly selected as an independent variable while the other was designated as a dependent variable.

(2) The model was fitted to the sample formed in step 1.

(3) Steps 1–2 were repeated 2500 times to produce a distribution function for each parameter of the model. The number of randomizations was chosen based on convergence of the mean and standard deviation for all the parameters of the model. Model residuals were tested for homogeneity for 5 samples (randomly selected from 2500 randomized samples formed in step 1), using residual plots as recommended in [20].

(4) Steps 1–3 were repeated for all combinations of the model parameters. For each combination, significance of a parameter ($P < 0.05$) was tested using P-value defined as $P = \min\{k, 2500 - k\}/2500$, where k is the number of randomizations for which the given parameter was less than or equal to the tested value. For example, when testing model (4) with parameters a, b (i.e, parameter c set to 1) for significance of parameter b for the Timmins site, fitting the model to 2500 randomized samples resulted in 2218 and 282 estimates of parameter b greater than 1 and less than or equal to 1, respectively, thus producing a P-value of 0.1124. The conclusion was that for model (4) with parameters a and b only, parameter b was not significantly different from 1 for the Timmins site.

(5) The best combination of model parameters for a given site was selected based on the highest adjusted R^2 and smallest standard error of estimate (SEE) from all "significant" combinations tested in Step 4 (i.e., combinations for which every parameter was significant). Adjusted R^2 and SEE for each parameter combination were calculated as the mean of adjusted R^2 and SEEs, respectively, for 2 500 individual randomizations described in Steps 1–3.

For each model, a combination of parameters that was consistently among the best for all sites was selected as the best variant of this model and used in the comparison among sites. A randomization technique [11] was used to test whether the models produced were statistically different among sites (i.e., to test whether the same model can be applied to all 6 sites). For each model, the tests were performed as follows:

(6) For each tree from the calibration set, one of the two stem diameters (either DBH or D03) was randomly selected as an independent variable while the other was designated as the dependent variable.

(7) The model was fitted to each site separately, and then to data from all 6 sites pooled. The F-statistic was calculated using the ratio of the difference between the residual sum of squares for the reduced and full models to the residual sum of squares for the full model divided by the appropriate degrees of freedom as in a conventional F-test [20].

(8) Each tree from the data set produced in Step 6 was randomly assigned to one of the 6 sites; the sample size for each site was kept equal to the size of the original sample (30 trees per site). The model was fitted to each site separately, then to all the sites pooled, and an F-ratio was calculated as in Step 7.

(9) Step 8 was repeated 5 000 times to produce a distribution of F-statistics. The F-statistic calculated in Step 7 was compared with this distribution to calculate the P-value for accepting the hypothesis that the sites are not significantly different.

(10) Steps 6–9 were repeated 5 times to ensure that the F-test results were not an artifact of a given random assignment of diameters in Step 6.

Model parameters were estimated for each site separately for the models that were significantly different among sites, and for all sites pooled together for the models not significantly different among sites. Steps 1–3 were repeated for each model, and resulting distributions of parameter values were used to calculate the mean and standard error for each parameter, as well as adjusted R^2 and SEE as described in Step 5. Finally, fitted models were evaluated using validation sets (82 trees not used for model calibration and comparison). Again, using the logic of the randomization technique, stem diameters for each validation tree were randomly assigned as independent and dependent variables, and the following statistics were calculated for each randomization: average prediction error (cm), standard deviation of prediction error (cm), relative prediction error (%) calculated as the ratio of predicted minus observed value divided by observed value and multiplied by 100, and standard deviation of relative prediction error (%). As with model calibration, 2500 randomizations were performed to calculate the mean values of these statistics.

All programs for this analysis were written in Borland Pascal. Procedures for generating random numbers and parameter estimation were taken from [17].

3. RESULTS

Testing of parameter significance revealed that models (1), (3), and (4) were overparametrized, i.e., either at least one parameter in each model was not significantly different ($P > 0.05$) from either zero or one, depending on the model tested, or equivalent predictions (in terms of R^2 and SEE) were produced by a model with fewer parameters. A comparison of models with fewer parameters among the sites (Steps 1–5 in Materials and methods) resulted in the following model variants that consistently provided the best predictions:

$$D_2 = D_1 \sqrt{\frac{(a-h_1)(z_2^2 + bz_2^3)}{(a-h_2)(z_1^2 + bz_1^3)}}, \qquad (1a)$$

$$D_2 = D_1 \left(\frac{z_2}{z_1}\right)^a, \qquad (2)$$

$$D_2 = D_1 \left[\frac{\exp z_2 - 1}{\exp z_1 - 1}\right]^c. \qquad (4a)$$

Removing parameter b turned model (3) into (2), thus reducing the number of models from 4 to 3.

Comparisons of each of the reduced models among sites produced the following results. Model (1a) was not significantly different among the 6 sites, with P-values for 5 simulations (Step 10 in Materials and methods) ranging from 0.278 to 0.445. Models (2) and (4a), however, differed significantly among the sites, with P-values for all 5 simulations for both models being less than 0.001. Therefore, model parameters and statistics of data fit were estimated for model (1a) using pooled data from all 6 sites, while similar estimates for models (2) and (4a) were performed for each site separately (Tab. II).

Model validation results are summarized in Table III. An example of predictions by model (1a) for an individual randomization of pooled samples from all 6 sites are presented in Figures 1a and 1b for calibration and validation data sets, respectively. Each point in the figures represents an individual tree, while the diagonal line indicates where observed values equal predicted values. Stem profiles predicted by the three fitted models are presented in Figure 2. Since models (2) and (4a) differed significantly among the sites, Figure 2 shows a stem profile generated using the mean value of parameters a and c, averaged over all 6 sites.

4. DISCUSSION

The stem profile equations developed in our study all had fewer parameters than the original models. This corroborates our assumption that the stem profile of young trees may be adequately described by a relatively simple model. The assumption was based on the fact that young trees usually have more tapered stems and less pronounced stem profile segmentation than mature trees [9, 10]. Model fitting showed that one or two parameters were sufficient to describe the data from young aspen trees. Other more complex models were considered but not tested because of the young age of the trees and relative simplicity of the data used in this study.

All models tested produced comparable results when fitted to calibration data (Tab. II): the adjusted coefficient of determination (R^2_{adj}) for all the models/sites ranged from 0.869 to 0.995, while the standard error of estimate (SEE) ranged from 0.06 to 0.19 cm. This may have resulted partly from the data, which included only 2 stem diameter measurements per tree for the trees ranging from 5 to 22 years old.

Comparing models among sites revealed no significant difference for model (1a) ($P > 0.278$), i.e., the stem profile of trees from all 6 sites was adequately described by a model with one set of estimated parameters. Validation of model (1a) produced mean prediction errors and mean relative prediction errors equal to less than 0.001 cm and 0.09%, respectively; standard deviations of these errors were equal to 0.15 cm and 4.09%, respectively. These values were of the same magnitude as the errors/standard deviations for models (2), (4a), fitted separately to each site.

Validation of models (2) and (4a) for site 3 produced a mean relative prediction error with a noticeable bias from zero and a larger standard deviation than for other sites. This may be the result of differences among stem profiles predicted by the fitted models (Fig. 2). Indeed, the upper part of the stem profile predicted by model (2) and (4a) has a much more pronounced paraboloidal form than that predicted by model (1a). Thus, if an

105-05897/

Table II. Means and standard errors of parameter estimates, adjusted R^2, and standard error of estimate (SEE) for model calibration. Models (2) and (4a) were significantly different among sites and therefore were fitted separately for each site.

Model		Parameter (standard error in brackets)			R^2_{adj}	SEE
		a	*b*	*c*		
Model (1a)		45.266 (14.498)	–0.522 (0.024)	–	0.995	0.144
Model (2)	Site 1	0.660 (0.011)	–	–	0.939	0.101
	Site 2	0.523 (0.007)	–	–	0.975	0.056
	Site 3	0.827 (0.018)	–	–	0.874	0.191
	Site 4	0.713 (0.005)	–	–	0.966	0.093
	Site 5	0.755 (0.008)	–	–	0.964	0.143
	Site 6	0.537 (0.003)	–	–	0.988	0.184
Model (4a)	Site 1	–	–	0.487 (0.008)	0.939	0.101
	Site 2	–	–	0.392 (0.005)	0.975	0.055
	Site 3	–	–	0.652 (0.017)	0.869	0.195
	Site 4	–	–	0.504 (0.004)	0.966	0.093
	Site 5	–	–	0.532 (0.006)	0.965	0.142
	Site 6	–	–	0.373 (0.002)	0.988	0.184

Table III. Means and standard deviations of absolute and relative prediction errors of model validation. Models (2) and (4a) were significantly different among sites and therefore were validated separately for each site.

Model		Mean prediction error (cm)	Mean standard deviation of error (cm)	Mean relative prediction error (%)	Mean standard deviation of relative error (%)
Model (1a) All sites pooled		–0.001	0.148	–0.092	4.094
Model (2)	Site 1	0.01	0.126	–0.179	6.453
	Site 2	0.001	0.083	–0.103	4.418
	Site 3	0.017	0.207	–1.76	18.23
	Site 4	0	0.065	–0.032	1.998
	Site 5	0.003	0.228	–0.141	4.749
	Site 6	–0.002	0.172	–0.013	2.013
Model (4a)	Site 1	0.01	0.129	–0.196	6.714
	Site 2	0.001	0.082	–0.103	4.398
	Site 3	0.022	0.219	–2.423	21.981
	Site 4	0.001	0.064	–0.032	1.998
	Site 5	0.003	0.229	–0.142	4.771
	Site 6	0.002	0.174	–0.013	2.028

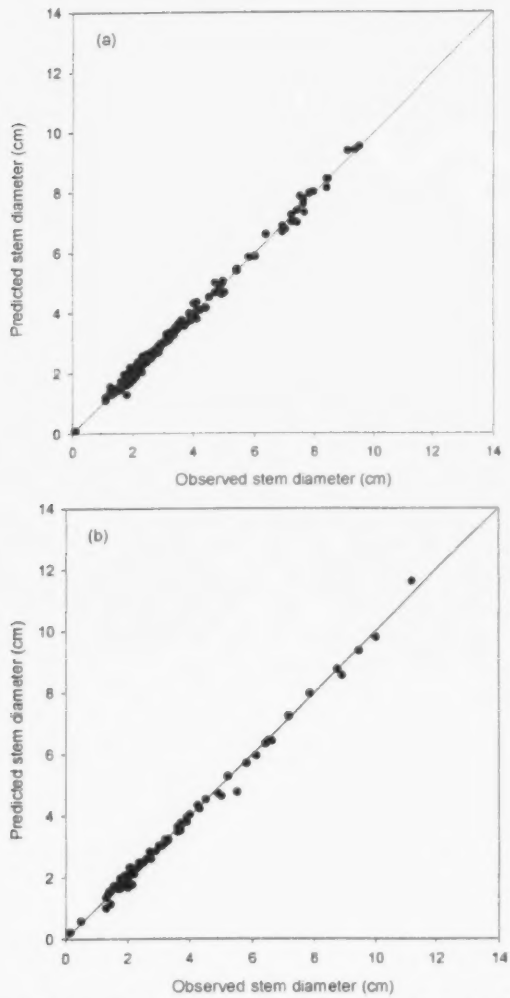

Figure 1. An example of predicted and observed stem diameters for an individual randomization of samples pooled from 6 sites. Predictions were generated by model (1a) for (a) calibration and (b) validation data.

input stem diameter is close to the top of the stem, models (2) and (4a) are likely to underestimate the diameter of the lower part of the stem compared to model (1a). Site 3 had the smallest trees, i.e., the highest percentage of all 6 sites where one of the two stem diameter measurements were taken close to the top of the stem. Therefore, relative prediction errors for models (2) and (4a) had a negative bias and larger standard deviation than for other sites (Tab. III).

As Table III shows, the means and standard deviations of both absolute and relative prediction errors produced by models (2) and (4a) were nearly identical when compared site by site, suggesting model similarity; see also Figure 2. Indeed, model (2) can be obtained by considering a Taylor series approximation of function (4a) in the neighbourhood of a point $z = 0$ and neglecting the terms beyond linear. Higher-order terms likely did not contribute substantially to the errors produced by the model (4a). Similarly, both original models (3) and (4) can be reduced to model (2) by considering only linear terms in the Taylor series approximation in the neighbourhood of point $z = 0$ and $az^b = 0$, respectively. However, models (2), (3), and (4) need to be tested with a larger data set to determine whether they are simply different algebraic approximations of the same family of stem profile curves.

Our approach produced a model applicable to all the sites in this study. Restricting the list of equations to only one preselected model (for example, any of models (2)–(4)) could have resulted in a need to include other site and tree variables in the model to explain the variation among the sites, as well as testing of various algebraic ways of incorporating these variables. This corroborates our initial hypothesis that testing of a broad list of models for significance of difference among sites is required prior to concluding that the trees belong to different populations.

The sites included in this study provided a range of stand age, stand density, and soil/climate conditions (Tab. I; see also [18]). Aspen stands on 3 sites (Timmins, Dryden, Mann Lake) were the same age (5 years old) but differed in stem density, basal area, and soil and climatic conditions. Four of the 6 study sites (Mann Lake, East Hillsport Lake, White Otter Lake, Flanders Road) were clustered together, and had very similar soil and climatic conditions, but the aspen stands ranged in age from

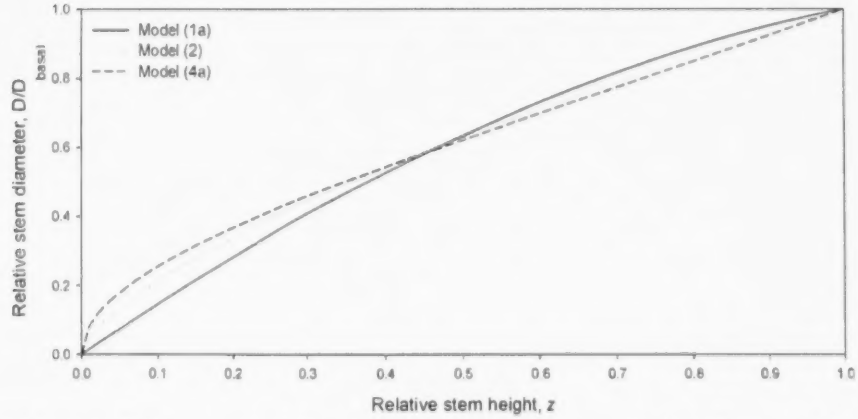

Figure 2. Relative stem diameter, D/D_{basal}, versus relative stem height, z, predicted by models (1a), (2), and (4a). Predictions by models (2) and (4a) used the mean value of parameters a and c, respectively, averaged over all 6 sites.

5 to 22. The study design allowed us to separate the effects of stand age, stem density, and soil/climatic conditions. Our results suggest that these variables may not be among the most important factors contributing to variation in stem profile equations within the studied range of these variables. Additional studies involving older stands and/or lower stem density are required to verify these results.

The use of bootstrap methods allowed us to alternate stem diameters measured for an individual tree as independent and dependent variables, thus providing an opportunity to extract 'maximum' information from the calibration data. With traditional statistics, inclusion of the same observation (tree) in the sample would have been impossible. Random assignment of diameters as dependent and independent variables can also be applied to studies when more than 2 measurements per tree are available. In the latter case, random subsampling of the data would help to avoid problems associated with correlation of measurements (and their errors) from the same tree. Similarly, using randomization as a method to compare an individual model among sites was free of assumptions about the error structure of the model, for example, the normal distribution of errors required in traditional *F*-tests. Here, the main underlying assumption was that the samples were representative of the general populations they came from. If those populations were the same, then random mixing of the samples would produce *F*-values of the same magnitude as that calculated for the original samples. On the contrary, if populations were different, then random mixing of the samples would result in *F*-values on average lower than that calculated for the original samples [11].

Equations (1a), (2), and (4a) can be used to predict the stem profile of young trembling aspen growing in conditions similar to those observed in this study. The study also sheds some light on the effect of several tree and site variables on the stem profile of young aspen. More importantly, we describe an approach that can be used to develop stem profile equations that are applicable to trees growing across a broad range of conditions.

Acknowledgements: We thank Jim Rice for data management; Doug Pitt for advice on data analysis and comments on the manuscript; Trudy Vaittinen for her help with illustrations; Lisa Buse for editing; and two anonymous reviewers for their helpful comments on a previous version of the manuscript.

REFERENCES

[1] Alemdag I.S., A ratio method for calculating stem volume to merchantable limits, and associated taper equations, For. Chron. 64 (1988) 18–26.

[2] Allen P.J., Average relative stem profile comparisons for three size classes of Caribbean pine, Can. J. For. Res. 23 (1993) 2594–2598.

[3] Crowley P.H., Resampling methods for computation-intensive data analysis in ecology and evolution, Annu. Rev. Ecol. Syst. 23 (1992) 405–447.

[4] Fang Z., Borders B.E., Bailey R.L., Compatible volume-taper models for loblolly and slash pine based on a system with segmented-stem form factors, For. Sci. 46 (2000) 1–12.

[5] Flewelling J.W., Raynes L.M., Variable-shape stem-profile predictions for western hemlock. Part I. Predictions from DBH and total height, Can. J. For. Res. 23 (1993) 520–536.

[6] Forslund R.R., The power function as a simple stem profile examination tool, Can. J. For. Res. 21 (1991) 193–198.

[7] Fowler J.H., Rennie J.C., Merchantable height in lieu of total height in stem profile equations, For. Sci. 34 (1988) 505–511.

[8] Kozak A., A variable exponent taper equation, Can. J. For. Res. 18 (1988) 1363–1368.

[9] Larson P.R., Stem form development of forest trees, For. Sci. Monogr. 5, 1963, 42 p.

[10] LeMay V.M., Kozak A., Marshall P.L., Muhairwe C., Literature review for development of a dynamic taper model for tree growth, Internal Report for the B.C. Science Council, Vancouver, British Columbia, 1991, 60 p.

[11] Manly B.F.J., Randomization, bootstrap and Monte Carlo methods in biology, Chapman & Hall, New York, 1997, 424 p.

[12] Morris D.M., Forslund R.R., The relative importance of competition, microsite, and climate in controlling the stem taper and profile shape in jack pine, Can. J. For. Res. 22 (1992) 1999–2003.

[13] Muhairwe C.K., LeMay V.M., Kozak A., Effects of adding tree, stand, and site variables to Kozak's variable-exponent taper equation, Can. J. For. Res. 24 (1994) 252–259.

[14] Newnham R.M., A variable-form taper equation, Canadian Forest Service, Petawawa National Forest Institute, Information Report PI-X-83, 1988.

[15] Ormerod D.W., A simple bole model, For. Chron. 49 (1973) 136–138.

[16] Pitt D.G., Kreutzweiser D.P., Application of computer-intensive statistical methods to environmental research, Ecotoxicol. Environ. Saf. 39 (1998) 78–97.

[17] Press W.H., Flannery B.P., Teukolsky S.A., Vetterling W.T., Numerical recipes in Pascal, Cambridge University Press, Port Chester, New York, 1994, 759 p.

[18] Rice J.A., MacDonald G.B., Weingartner D.H., Pre-commercial thinning of trembling aspen in northern Ontario: Part 1, Growth responses, For. Chron. 77 (2001) 893–901.

[19] Rustagi K.P., Loveless R.S., Compatible variable-form volume and stem profile equations for Douglas-fir, Can. J. For. Res. 21 (1991) 143–151.

[20] Seber G.A.F., Linear Regression Analysis, John Wiley and Sons, New York, 1977, 465 p.

[21] Valentine H.T., Gregoire T.G., A switching model of bole taper, Can. J. For. Res. 31 (2001) 1400–1409.

[22] Zakrzewski W.T., A mathematically tractable stem profile model for jack pine in Ontario, North. J. Appl. For. 16 (1999) 138–143.